

Reservoir Computing with Signal-Mixing Cavities

Floris Laporte¹, Joni Dambre², and Peter Bienstman¹

¹Photonics Research Group, UGent – imec, Technologiepark 15, 9052 Gent, Belgium

²IDLab, UGent - imec, Technologiepark 15, 9052 Gent, Belgium

Tel: +32 9 264 3450, Fax: +32 9 264 3593, e-mail: floris.laporte@ugent.be

ABSTRACT

In an age where we get swamped by big data, new machine learning techniques for efficient high-speed data processing become more important than ever. One of these techniques, known as reservoir computing, is specifically designed for processing time-dependent data. We propose some new ideas for implementing such a reservoir computer on a silicon photonics chip for low-power and high-bandwidth optical communication applications. Our simulations show that this photonic reservoir can for example be used in pattern recognition tasks such as header recognition.

Keywords: optical computing, photonic reservoir computing, pattern recognition, photonic crystal cavities.

1. INTRODUCTION

Reservoir Computing is a well-established machine learning branch focusing on the processing of temporal data. It was first proposed in the early 2000s [1,2] as a way of getting around the difficulties of training a neural network with internal feedback, which is necessary for processing time-dependent signals. In reservoir computing, one uses a highly dynamical and usually nonlinear system as a signal-mixing *reservoir*, after which a large quantity of operations can be performed by using a suitable readout which performs an optimized linear combination of the internal reservoir states (Fig. 1).

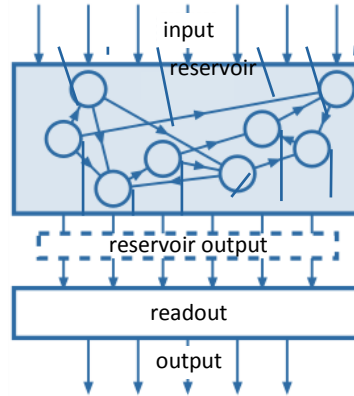


Figure 1. Typical Reservoir operation scheme. A time-dependent and possibly spatially varying input is fed into a recurrent neural network. In this so-called reservoir, the input signal mixes with itself. This mixed signal is in turn read out at fixed time intervals by a specifically trained readout, which makes an application dependent linear combination of the reservoir states.

Although reservoir computing was originally proposed as a software algorithm, any sufficiently high-dimensional dynamical and nonlinear system can act as a reservoir [3]. Since the computationally hardest tasks are done by the reservoir (the readout just performs a linear combination of the reservoir states), the straightforward next step in creating an efficient reservoir computer is to look for a suitable hardware implementation. Efficient mechanical [4], electronic [5,6] and photonic [7] solutions have been implemented in practice.

Photonics has a few advantages over the other implementations, in the sense that it is the perfect choice in the context of telecom applications. First, it enables signal processing at the speed of light, while ideally cutting out the middle step of converting the signal into an electronic signal and back. Second, it provides massive parallelism inherent to optical structures. Our previous implementations of a photonic reservoir have so far shown promise [8,9], but these structures usually remain very faithful to the recurrent node structure of the echo state network reservoir [1], originally proposed in the context of neural networks, which can limit their power efficiency and results in devices with a large footprint.

This paper proposes instead a simple photonic crystal cavity (Fig. 2) with a quarter stadium shape [10], that is known to foster interesting mixing dynamics [6,7], while also exhibiting a form of memory, as the signal is bound to remain in the cavity for a certain amount of time depending on its Q-factor. In the context of reservoir computing, exhibiting such a *fading memory* together with the *nonlinearity* of the system turn out to be the key properties for a successful operation [3]. Since detecting the optical signals coming out of the attached cavity

waveguides is an inherently quadratic nonlinear effect [8], the proposed system should be able to act as a decent reservoir, while leaving the conventional node-structure behind.

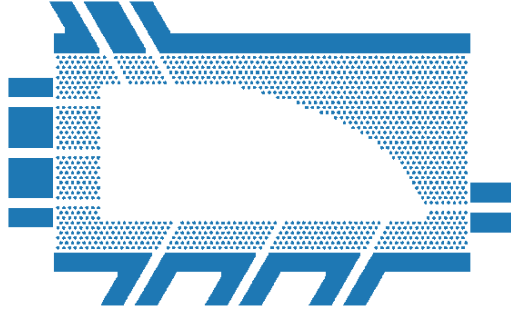


Figure 2. Photonic crystal cavity with a quarter stadium shape used as a reservoir; blue:oxide/holes, white:silicon

2. DESIGN

The reservoir consists of a $30\mu\text{m} \times 15\mu\text{m}$ photonic crystal cavity with a quarter-stadium shape with several attached waveguides. Any of those attached waveguides can be used as either an input or an output waveguide. The dimensions of the cavity should match the time scale of the signal to be processed: to perform operations at a certain bitrate, the lifetime of the cavity should be at least a multiple of the bit period.

To increase the lifetime of the cavity, one needs to thoroughly optimize the Q-factor, and reduce the number of attached waveguides to a bare minimum. Additionally, since the propagation losses in the cavity are low, one can also increase the size of the cavity at the cost of a less rich mixing of the fields. It is thus a non-trivial task to optimize the cavity properties without sacrificing computing power. This trade-off will also depend a lot on the application being studied. However, for the sake of consistency throughout this paper, we chose for the cavity with 9 attached waveguides depicted in Fig. 2 exhibiting a Q-factor of about 5000, which yields a cavity half-life $T_{1/2} = 5$ ps.

While the shape of the cavity introduces interesting dynamics of the fields, it must be stressed that this is not a truly chaotic phenomenon, since the state of the field does in the limit not depend on its initial conditions. This makes it behave exactly like a fading memory necessary for reservoir computing.

3. SIMULATIONS

The simulation of the photonic crystal was done using Lumerical FDTD simulations. A single pulse was sent through the crystal after which the pulse response (Fig. 3) was recorded. After this, the pulse responses were coherently added together to form a bit stream response, which was then fed into the detector model. These steps were repeated for every desired bitrate (pulse length). These detected signals were subsequently fed into the machine learning model which was specifically trained to perform a certain task. The two tasks considered here were a benchmark XOR task and a header recognition task.

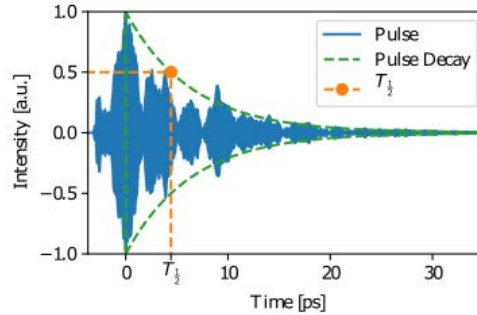


Figure 3. The pulse response of a 3 ps pulse at one of the exit waveguides. An introduced pulse needs a certain time to decay, characterized by $T_{1/2}$, which is the characteristic timescale of the cavity.

In the detector, the complex signal consisting of an amplitude and phase results in a real-valued magnitude of the signal. However, what happens at the detector is a lot more than just this quadratic nonlinearity. Since the time scales of the tiny variations in the pulse response depicted in Fig. 3 are a lot smaller than what the detector can resolve, we have to take into account the limited bandwidth and other physical properties of the detector.

In our simulations, we assumed a high-impedance ($R = 100\Omega$) detector with a bandwidth $\Delta f = 100\text{GHz}$ and a responsivity $\eta = 1\text{A/W}$. The white noise I_{tn} was modelled as thermal Nyquist noise [13], while shot noise I_{sn} was introduced by assuming a Poisson process [14], yielding for the total noise of the detector

$$I_n = \sqrt{I_{tn}^2 + I_{sn}^2} = \sqrt{\left(\frac{4kT\Delta f}{R}\right)^2 + (2qI\Delta f)^2}$$

This noise formula yields a maximal noise amplitude of about 6μA for an input intensity of 1mW. This is mostly carried by the thermal noise, which yields a signal-to-noise ratio (SNR) of about 10 for some of the exits (Fig. 4). In turn, the detected signal gets sampled at fixed time intervals (Fig. 4) to be combined into a single output signal using pre-trained weights that are specifically optimized for a certain application.

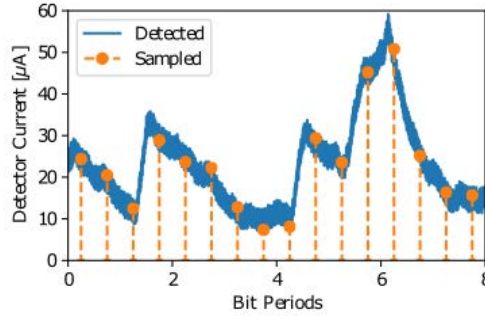


Figure 4. The detected signal with average intensity of 1mW gets sampled at discrete time intervals equal to or a fraction of the bit period of the signal.

4. RESULTS

4.1 XOR Memory on neighboring bits

A good benchmark for the ability to perform Boolean operations is the nonlinear XOR task, where the XOR is taken between two bits b_n and b_{n-k} , k bits apart. Since a normal conventional linear classifier can only achieve a minimum of 25% error rate, it is also a good performance indicator of the nonlinearity in the system. In Fig. 5, the performance is shown for neighboring bits ($k = 1$) and two bits with one bit in between ($k = 2$). A good performance indicator is the amount of time this operation can be performed after receiving the last bit. Obviously, the larger this region of operation, the more stable the task will be. In Fig. 5, we see a clear region of operation where the XOR of neighboring bits can be found without error, while the region of operation for the XOR between two bits with one bit in between is a lot smaller.

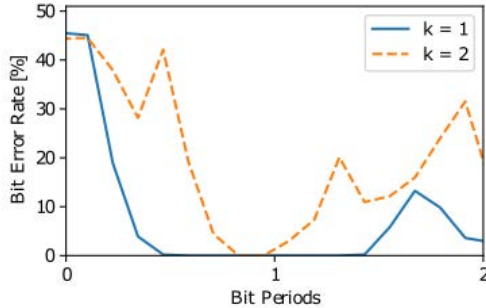


Figure 5. Error rate in function of the reproduction delay for different bit separation k at 100Gbps.

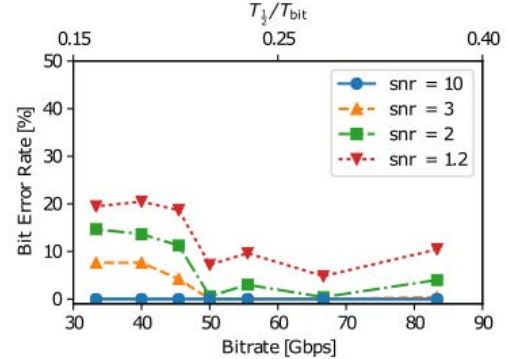


Figure 6. The optimal bitrate for the XOR task lies at 50 Gbps. For this bitrate, we are able to get good performance at a SNR of down to a minimum of 2, while there is a full band of good performance for a SNR of 3 at higher bitrates.

However, probably a more interesting measure of stability, is the range of bitrates for which this task is performed optimally by the reservoir. Figure 6 shows that the performance of the reservoir for the XOR operation between neighboring bits is optimal at about 50 Gbps. This bitrate might seem high, but this optimal bitrate can be easily reduced by creating a bigger cavity. Therefore, we also choose to quantify the optimal performance by the dimensionless parameter $T_{1/2}/T_{\text{bit}}$. This allows us to easily scale the reservoir operating range by tweaking its design parameters.

4.2 Header Recognition

As a more useful and general task, the performance of the reservoir in recognizing headers in a bit stream was assessed. The simple cavity performs reasonably well for header recognition tasks, with an errorless recognition and a large region of operation of up to 4-bit headers. Note that this performance is really good, as a cascaded network should be able to vastly outperform this network by training each of the cascaded systems for a different part of the header.

In this case, Linear Discriminant Analysis was used as the linear classification algorithm. This algorithm allows us to project the 9-dimensional (# connected waveguides) reservoir output state to a lower dimensional state.

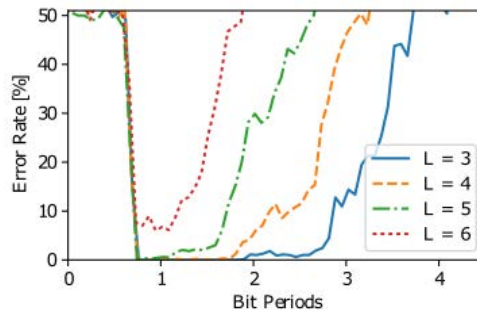


Figure 7. By sweeping over the delay to find the operation range, we find that the reservoir can distinguish headers of up to $L = 4$ bits without error at a bit rate of 100 Gbps.

5. FUTURE WORK

The next step of this project is to create an on-chip version of the proposed work.

A second challenge would be to create an integrated implementation of the readout in the optical domain. Such a readout would allow us to perform the proposed classification problems quasi-instantaneously while being more power- and cost-efficient than using a separate detector for each channel. This would however come at the cost of the trainability of the system, which is why good accurate simulations can still yield valuable results.

A third task is to investigate the scalability to larger reservoirs, as is required for more complex tasks. First simulations show that the reservoir operation can already be improved by just adding a second delayed input.

6. CONCLUSIONS

Photonic cavities on chip seem to be the ideal candidate for an optical reservoir computer. The performance on the nonlinear Boolean XOR task is good while the performance on header recognition tasks. Performance on these memory-dependent problems can probably be increased by optimizing the Q-factor even further or by increasing the size of the cavity. Additionally, the computing power of these reservoirs can most likely still be increased by not solely relying on the nonlinearity of the detector, but also by adding extra nonlinear materials in the cavity or on top of the silicon bulk.

ACKNOWLEDGEMENTS

Part of this work was performed in the context of the Belgian IAP program Photonics@BE, as well as the European H2020 project Phresco.

REFERENCES

- [1] H. Jaeger, "The echo state approach to analysing and training recurrent neural networks," *GMD Rep.*, vol. 148, no. 34, p. 13, 2001.
- [2] W. Maass, T. Natschl ger, and H. Markram, "Real-time computing without stable states: A new framework for neural computation based on perturbations," *Neural Comput.*, 2002.
- [3] B. Schrauwen, D. Verstraeten, and J. Van Campenhout, "An overview of reservoir computing: Theory, applications and implementations," in *Proc. 15th Eur. Symp. Artif. Neural Networks*, pp. 471–82, Apr. 2007.
- [4] H. Hauser, A. Ijspeert, R. F uchslin, R. Pfeifer, and W. Maass, "Towards a theoretical foundation for morphological computation with compliant bodies," *Biol. Cybern.*, vol. 105, no. 2011, pp. 355–370, 2011.
- [5] N. D. Haynes, M. C. Soriano, D. P. Rosin, I. Fischer, and D. J. Gauthier, "Reservoir computing with a single time-delay autonomous Boolean node," *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, vol. 91, no. 2, pp. 1–6, 2015.
- [6] L. Appeltant *et al.*, "Information processing using a single dynamical node as complex system," *Nat. Commun.*, vol. 2, pp. 466–468, 2011.
- [7] Y. Paquot *et al.*, "Optoelectronic reservoir computing," *Sci. Rep.*, vol. 2, p. 287, 2012.
- [8] K. Vandoorne *et al.*, "Experimental demonstration of reservoir computing on a silicon photonics chip," *Nat. Commun.*, vol. 5, p. 3541, 2014.
- [9] K. Vandoorne, M. Fiers, B. Schrauwen, J. Dambre, and P. Bienstman, "Photonic reservoir computing: A New approach to optical information processing."
- [10] S. D. Cohen, H. L. D. de S. Cavalcante, and D. J. Gauthier, "Subwavelength position sensing using nonlinear feedback and wave chaos," *Phys. Rev. Lett.*, vol. 107, no. 25, p. 254103, 2011.
- [11] H. Stockmann and J. Stein, "Quantum Chaos in billiards studied by microwave absorption," *Phys. Rev. Lett.*, vol. 64, no. 19, p. 2215, 1990.
- [12] C. Liu *et al.*, "Triggering extreme events at the nanoscale in photonic seas," *Nat. Phys.*, vol. 11, no. 4, pp. 358–363, 2015.
- [13] H. Nyquist, "Thermal agitation of electric charge in conductors," *Phys. Rev.*, vol. 32, no. 1, pp. 110–113, 1928.
- [14] Y. M. Blanter and M. Buttiker, "Shot noise in mesoscopic conductors," *Phys. Rep.*, vol. 336, no. 1–2, pp. 1–166, 2000.